lines is surely possible with the photocopier that is ChatGPT, so, in that sense, the answer is yes. But I don't think that anyone would claim that photocopiers have become an essential tool in the creation of art; the vast majority of artists don't use them in their creative process, and no one argues that they're putting themselves at a disadvantage with that choice.

So let's assume that we're not talking about a new genre of writing that's analogous to Xerox art. Given that stipulation, can the text generated by large language models be a useful starting point for writers to build off when writing something original, whether it's fiction or nonfiction? Will letting a large language model handle the boilerplate allow writers to focus their attention on the really creative parts?

Obviously, no one can speak for all writers, but let me make the argument that starting with a blurry copy of unoriginal work isn't a good way to create original work. If you're a writer, you will write a lot of unoriginal work before you write something original. And the time and effort expended on that unoriginal work isn't wasted; on the contrary, I would suggest that it is precisely what enables you to eventually create something original. The hours spent choosing the right word and rearranging sentences to better follow one another are what teach you how meaning is conveyed by prose. Having students write essays isn't merely a way to test their grasp of the material; it gives them experience in articulating their thoughts. If students never have to write essays that we have all read before, they will never gain the skills needed to write something that we have never read. And it's not the case that, once you have ceased to be a student, you can safely use the template that a large language model provides. The struggle to express your thoughts doesn't disappear once you graduate—it can take place every time you start drafting a new piece. Sometimes it's only in the process of writing that you discover your original ideas. Some might say that the output of large language models doesn't look all that different from a human writer's first draft, but, again, I think this is a superficial resemblance. Your first draft isn't an unoriginal idea

expressed clearly; it's an original idea expressed poorly, and it is accompanied by your amorphous dissatisfaction, your awareness of the distance between what it says and what you want it to say. That's what directs you during rewriting, and that's one of the things lacking when you start with text generated by an A.I.

There's nothing magical or mystical about writing, but it involves more than placing an existing document on an unreliable photocopier and pressing the Print button. It's possible that, in the future, we will build an A.I. that is capable of writing good prose based on nothing but its own experience of the world. The day we achieve that will be momentous indeed—but that day lies far beyond our prediction horizon. In the meantime, it's reasonable to ask, What use is there in having something that rephrases the Web? If we were losing our access to the Internet forever and had to store a copy on a private server with limited space, a large language model like ChatGPT might be a good solution, assuming that it could be kept from fabricating. But we aren't losing our access to the Internet. So just how much use is a blurry JPEG, when you still have the original? ♦

MORE SCIENCE AND TECHNOLOGY

- A security camera <u>for the entire planet</u>.

- In India, a clean-power plant the size of Manhattan could be a model for the world—<u>or a cautionary tale</u>.

- Saving the climate will depend on blue-collar workers. Can we train enough of them <u>before time runs out</u>?

- Cargo ships are among the dirtiest vehicles in existence. Can an ancient technology <u>clean them up</u>?

- <u>A case for climate optimism</u>, and pragmatism, from John Podesta.

- A heat shield for <u>the most important ice on Earth</u>.

11

- The climate solutions <u>we can't live without</u>.

<u>Sign up</u> for our daily newsletter to receive the best stories from *The New Yorker*.

<u>*Ted Chiang*</u> *is an award-winning author of science fiction. In 2016, the title story from his first collection, "*<u>Stories of Your Life and Others</u>," was adapted into the film "Arrival." He lives in Bellevue, Washington, where he works as a freelance technical writer.*

More: **Artificial Intelligence (A.I.)**   **Internet**   **Writing**   **Images**   **Algorithms**   **Technology**

THE DAILY

The best of *The New Yorker*, every day, in your in-box, plus occasional alerts when we publish major stories.

E-mail address

> Your e-mail address

> **Sign up**

By signing up, you agree to our <u>User Agreement</u> and <u>Privacy Policy & Cookie Statement</u>.

Read More

Renew WIRED + Send a Gift   RENEW + GIVE

DAVID NIELD   BUSINESS   APR 30, 2023 7:00 AM

How ChatGPT and Other LLMs Work—and Where They Could Go Next

Large language models like AI chatbots seem to be everywhere. If you understand them better, you can use them better.
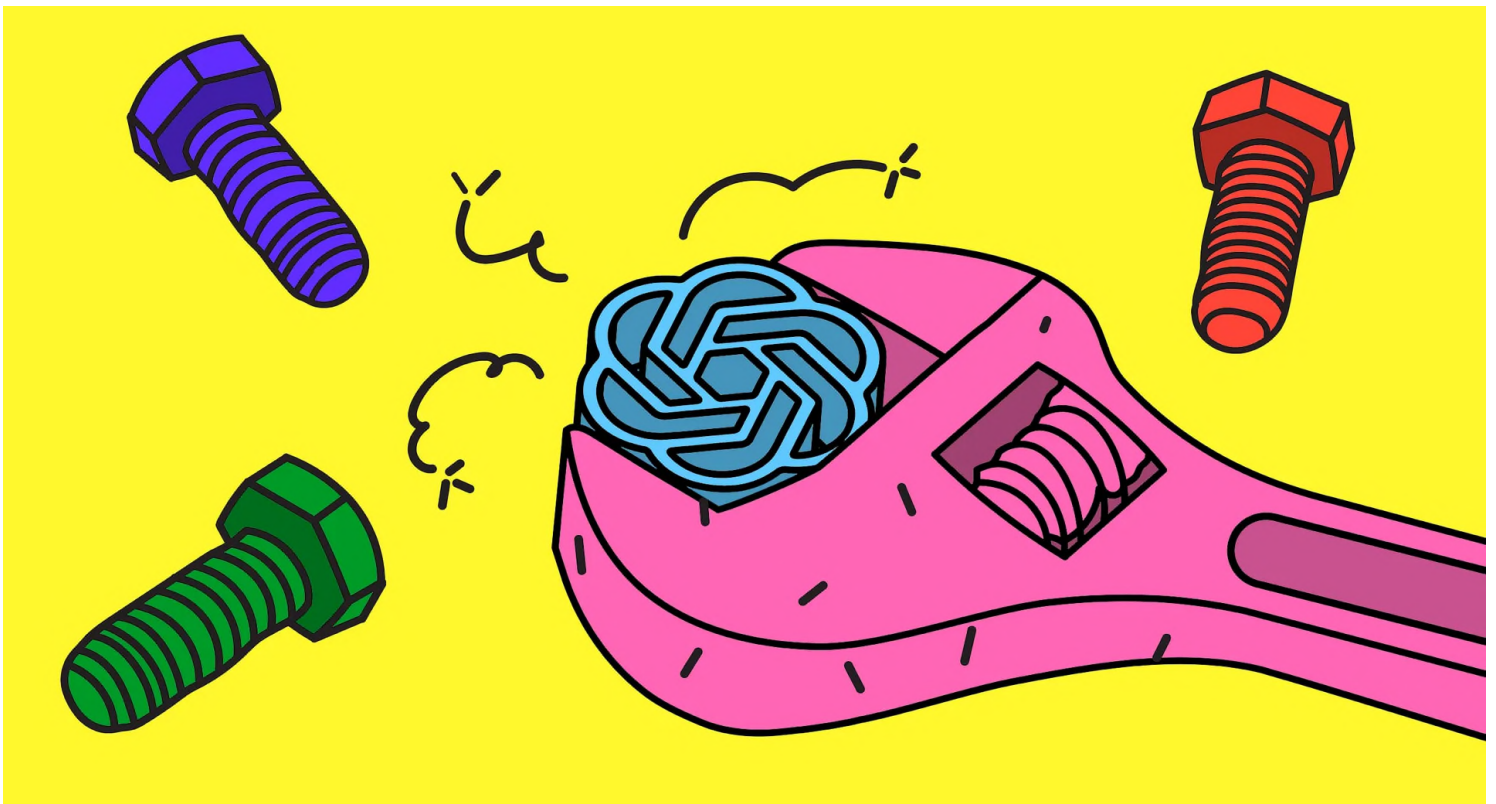

ILLUSTRATION: JAMES MARSHALL; GETTY IMAGES

13

AI-POWERED CHATBOTS SUCH as <u>ChatGPT</u> and <u>Google Bard</u> are certainly having a moment—the next generation of conversational software tools promise to do everything from taking over our web searches to producing an endless supply of creative literature to remembering all the world's knowledge so we don't have to.

ChatGPT, Google Bard, and other bots like them, are examples of <u>large language models</u>, or LLMs, and it's worth digging into how they work. It means you'll be able to better make use of them, and have a better appreciation of what they're good at (and what they really shouldn't be trusted with).

What's New With WIRED

How ChatGPT and Other LLMs Work and Whe

00:00 / 06:42

Like a lot of artificial intelligence systems—like the ones designed to recognize your voice or generate cat pictures—LLMs are trained on huge amounts of data. The companies behind them have been rather circumspect when it comes to revealing where exactly that data comes from, but there are certain clues we can look at.

For example, <u>the research paper</u> introducing the LaMDA (Language Model for Dialogue Applications) model, which Bard is built on, mentions Wikipedia, "public forums," and "code documents from sites related to programming like Q&A sites, tutorials, etc." Meanwhile, Reddit <u>wants to start charging</u> for access to its 18 years of text conversations, and <u>StackOverflow just announced</u> plans to start charging as well. The implication here is that LLMs have been making extensive use of both sites up until this point as sources, entirely for free and on the backs of the people who built and used those resources. It's clear that a lot of what's publicly available on the web has been scraped and analyzed by LLMs.

14

LLMs use a combination of machine learning and human input. OPENAI VIA DAVID NIELD

All of this text data, wherever it comes from, is processed through a neural network, a commonly used type of AI engine made up of multiple nodes and layers. These networks continually adjust the way they interpret and make sense of data based on a host of factors, including the results of previous trial and error. Most LLMs use a specific neural network architecture called a transformer, which has some tricks particularly suited to language processing. (That GPT after Chat stands for Generative Pretrained Transformer.)

Specifically, a transformer can read vast amounts of text, spot patterns in how words and phrases relate to each other, and then make predictions about what words should come next. You may have heard LLMs being compared to supercharged autocorrect engines, and that's actually not too far off the mark: ChatGPT and Bard don't really "know" anything, but they are very good at figuring out which word follows another, which starts to look like real thought and creativity when it gets to an advanced enough stage.

15

One of the key innovations of these transformers is the self-attention mechanism. It's difficult to explain in a paragraph, but in essence it means words in a sentence aren't considered in isolation, but also in relation to each other in a variety of sophisticated ways. It allows for a greater level of comprehension than would otherwise be possible.

There is some randomness and variation built into the code, which is why you won't get the same response from a transformer chatbot every time. This autocorrect idea also explains how errors can creep in. On a fundamental level, ChatGPT and Google Bard don't know what's accurate and what isn't. They're looking for responses that seem plausible and natural, and that match up with the data they've been trained on.

So, for example, a bot might not always choose the most likely word that comes next, but the second- or third-most likely. Push this too far, though, and the sentences stop making sense, which is why LLMs are in a constant state of self-analysis and self-correction. Part of a response is of course down to the input, which is why you can ask these chatbots to simplify their responses or make them more complex.

You might also notice generated text being rather generic or clichéd—perhaps to be expected from a chatbot that's trying to synthesize responses from giant repositories of existing text. In some ways these bots are churning out sentences in the same way that a spreadsheet tries to find the average of a group of numbers, leaving you with output that's completely unremarkable and middle-of-the-road. Get ChatGPT to talk like a cowboy, for instance, and it'll be the most unsubtle and obvious cowboy possible.

Human beings are involved in all of this too (so we're not quite redundant, yet): Trained supervisors and end users alike help to train LLMs by pointing out mistakes, ranking answers based on how good they are, and giving the AI high-quality results to aim for. Technically, it's known as "reinforcement learning on human feedback" (RLHF). LLMs then refine their internal neural networks further to get better results next time. (These are still relatively early days for the technology at this level, but we've already seen numerous notices of upgrades and improvements from developers.)

As these LLMs get bigger and more complex, their capabilities will improve. We know that ChatGPT-4 has in the region of 1 trillion parameters (although OpenAI won't confirm,) up from 175 billion in ChatGPT 3.5—a parameter being a mathematical relationship linking words through numbers and algorithms. That's a vast leap in terms of understanding relationships between words and knowing how to stitch them together to create a response.

See What's Next in Tech With the Fast Forward Newsletter

A weekly dispatch from the future by Will Knight, exploring AI advances and other technology set to change our lives. Delivered every Thursday.

By signing up you agree to our User Agreement (including the class action waiver and arbitration provisions), our Privacy Policy & Cookie Statement and to receive marketing and account-related emails from WIRED. You can unsubscribe at any time.

From the way LLMs work, it's clear that they're excellent at mimicking text they've been trained on, and producing text that sounds natural and informed, albeit a little

bland. Through their "advanced autocorrect" method, they're going to get facts right most of the time. (It's clear what follows "the first president of the USA was …") But it's here where they can start to fall down: The most *likely* next word isn't always the *right* one.

Correction, 5/9/2023: *A previous version of this story underestimated how many parameters ChatGPT 3.5 had (it's 175 billion, not 175 million) and stated ChatGPT 4 had upwards of 100 trillion, but reporting between the time this story was published and now indicates the true number may be as low as 1 trillion.*

Get More From WIRED

- 📧 Get the best stories from <u>WIRED's iconic archive</u> in your inbox
- 🎧 Our new <u>podcast</u> wants you to <u>*Have a Nice Future*</u>
- The explosive legacy of the <u>pandemic hand sanitizer boom</u>
- Scientists <u>gave people psychedelics</u>—then erased their memory
- I asked <u>AI chatbots to help me shop</u>. They all failed
- The race is on to crack an artist's <u>"test" signal from aliens</u>
- The speedrunners trying to break <u>*Tears of the Kingdom*</u>
- 🏔️ Embrace the new season with the Gear team's best picks for best <u>tents</u>, <u>umbrellas</u>, and <u>robot vacuums</u>

<u>David Nield</u> is a tech journalist from Manchester in the UK, who has been writing about apps and gadgets for more than two decades. You can follow him on <u>Twitter</u>.

CONTRIBUTOR   🐦

TOPICS   ARTIFICIAL INTELLIGENCE   CHATBOT   CHATGPT

MORE FROM WIRED

Renew WIRED + Send a Gift   RENEW + GIVE

WILL KNIGHT   BUSINESS   MAY 11, 2023 12:00 PM

The Boring Future of Generative AI

ChatGPT's chaotic streak can be charming. Google's new chat-style search shows text-generation technology is headed in a much tamer direction.